Tensor Encoding / Decoding:

Derek K Jones
Centre for Neuroimaging Sciences
Institute of Psychiatry, P089, London, UK.
Email: d.jones@iop.kcl.ac.uk

INTRODUCTION

In the first talk of the session, Dr Peter Basser will have introduced the theory of diffusion in biological systems and shown how, in tissue with ordered structure on the length scale of the measurement volume (i.e., voxel), diffusion appears anisotropic. In such cases, the diffusion properties cannot be fully characterized by a scalar value. The next most complex model of diffusion is the diffusion tensor – a 3 x 3 symmetrical positive definite matrix. In this presentation, we will focus on how the diffusion is estimated and how quantitative parameters are derived from the tensor, with focus on obtaining robust values. Pitfalls, artefacts and their remediation will be covered in the talk by Dr Carlo Pierpaoli, while sequences will be covered by Dr Jim Pipe.

The Basics

Diffusion Weighted MR: In essence, the diffusion sensitivity of most MR pulse sequences can be increased by adding a pair of pulsed magnetic field gradients into the sequence. The effect of the gradients is to introduce a spatially dependent phase accumulation. The sequence is designed so that, for a stationary spin, the phase accumulation due to the first gradient is matched in amplitude but **reversed in sign** to the phase accumulation due to the second gradient. The result is that for stationary spins, the net phase change is zero. For spins that move during the experiment (i.e., diffusing water molecules), the phase accumulations due to the 1st and 2nd gradients are not matched in amplitude – resulting in a net phase change. As diffusion is a random walk, the motion of spins is *incoherent* within the voxel – which results in a *distribution* of phases – resulting in a loss of phase coherence and signal attenuation. The amount of diffusion weighting will depend on the strength of the encoding gradients (i.e., how strongly the phase accumulation depends on spatial location), and on the duration of the experiment (i.e., for how long diffusional motion and phase dispersion is allowed to occur). These parameters are summarised into a single number called the '*b*-factor'. For a simple pair of pulsed gradients, the b-factor is given by the so-called Stejskal-Tanner expression:

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3} \right), \qquad [1]$$

where γ is the gyromagnetic ratio, G is the gradient amplitude, and δ and Δ are the temporal duration and separation of the diffusion-encoding gradients, respectively.

Diffusion Weighted Signal in Isotropic Media:

For isotropic media, the signal attenuation is written as:

$$I = I_0 \exp(-bD), \qquad [2]$$

where I and I_0 are the diffusion-weighted and non-diffusion-weighted intensities, respectively, b is the b-factor and D is the scalar apparent diffusion coefficient. It should be clear that by acquiring an additional image with no diffusion weighting (i.e. I_0), and with knowledge of the b-factor, it is possible to estimate the diffusion coefficient directly.

Diffusion-Weighted Signal in Anisotropic Media

As discussed in the talks by Drs Basser and Beaulieu, diffusion in anisotropic systems is more completely characterized by a tensor matrix. In such situations, Eq. [1] has to be re-written by replacing the scalar b-factor with a b-matrix, whose elements b_{ij}, scale the attenuation of the signal by the corresponding elements of the diffusion tensor, D_{ij}. Thus, Eq. [1] becomes:

$$I = I_0 \exp\left(-b_{xx}D_{xx} - b_{yy}D_{yy} - b_{zz}D_{zz} - 2b_{xy}D_{xy} - 2b_{xz}D_{xz} - 2b_{yz}D_{yz}\right), \qquad [3]$$

For a detailed description of the computation of the b-matrix the interested reader is referred to Mattiello *et al*. (1994), but the key point is that the elements of the b-matrix can be varied by changing the direction of the diffusion encoding gradient. If the i^{th} encoding gradient has components, $\mathbf{g}^i = \left[g_x^i \; g_y^i \; g_z^i\right]$, and b is the trace of the b-matrix, then

$$\left[b_{xx}^i, b_{yy}^i, b_{zz}^i, b_{xy}^i, b_{xz}^i, b_{yz}^i\right] = b\left[\left(g_x^i\right)^2, \left(g_y^i\right)^2, \left(g_z^i\right)^2, \left(g_x^i g_y^i\right), \left(g_x^i g_z^i\right), \left(g_y^i g_z^i\right)\right], \qquad [4]$$

Estimating the Diffusion Tensor

In linear algebra, when one has n unknowns, one needs a set of n simultaneous equations to be able to solve for those unknowns. Since the diffusion tensor is a 3 x 3 symmetrical matrix, it contains just six unique elements – or six 'unknowns'. Therefore, to estimate the diffusion tensor, it is sufficient to obtain a set of six simultaneous linear equations, containing the elements, to estimate the full tensor. To obtain a linear equation, a logarithmic transform of each side of Eq. [2] is performed. To set up simultaneous equations, the direction of the applied encoding gradients is varied – in at least six non-collinear and non-coplanar directions. If only six measurements are made, then the tensor can be found by simple matrix inversion and the solution is exact.

If $$\mathbf{B} = \begin{bmatrix} b_{xx}^1 & b_{yy}^1 & b_{zz}^1 & 2b_{xy}^1 & 2b_{xz}^1 & 2b_{yz}^1 \\ b_{xx}^2 & b_{yy}^2 & b_{zz}^2 & 2b_{xy}^2 & 2b_{xz}^2 & 2b_{yz}^2 \\ b_{xx}^3 & b_{yy}^3 & b_{zz}^3 & 2b_{xy}^3 & 2b_{xz}^3 & 2b_{yz}^3 \\ b_{xx}^4 & b_{yy}^4 & b_{zz}^4 & 2b_{xy}^4 & 2b_{xz}^4 & 2b_{yz}^4 \\ b_{xx}^5 & b_{yy}^5 & b_{zz}^5 & 2b_{xy}^5 & 2b_{xz}^5 & 2b_{yz}^5 \\ b_{xx}^6 & b_{yy}^6 & b_{zz}^6 & 2b_{xy}^6 & 2b_{xz}^6 & 2b_{yz}^6 \end{bmatrix}, \mathbf{S} = \begin{bmatrix} \ln(I_0/I_1) \\ \ln(I_0/I_2) \\ \ln(I_0/I_3) \\ \ln(I_0/I_4) \\ \ln(I_0/I_5) \\ \ln(I_0/I_6) \end{bmatrix} \text{ and } \mathbf{D} = \begin{bmatrix} D_{xx} \\ D_{yy} \\ D_{zz} \\ D_{xy} \\ D_{xz} \\ D_{yz} \end{bmatrix},$$

then we can write $\mathbf{S} = \mathbf{BD}$, which is readily solved for $\mathbf{D}$, i.e., $\mathbf{D} = \mathbf{B}^{-1}\mathbf{S}$. (It is straightforward to find the inverse of $\mathbf{B}$ in this case, since it is a square matrix).

In the example just given, the model fits the data points exactly and therefore includes the noise contamination of the data. Unless scanning in the absolute minimum time possible, then this is far from ideal – and more samples of the diffusion-weighted signal should be obtained. The tensor is then estimated using regression. There are two main approaches: The first is to use non-linear optimization routines (e.g. Levenberg-Marquardt) to estimate the elements of the tensor directly from the diffusion-weighted signals, which are expressed as in Eq. [3].

This has the advantage that the errors in the signals are *homoscedastic*. However, non-linear fitting routines: (a) generally require an initial estimate; (b) Require the operator to choose a convergence limit; (c) Often take a long time to converge; (d) Are prone to local minima. On the basis of these limitations, the linear multivariate regression approach (Basser *et al.* 1994) is favoured. With more than six measurements, the matrix **B** above is no longer square, and so we have to compute the *pseudo*-inverse. Further, the logarithmic transformation of the data introduces *heteroscedasticity* in the errors. Each data point should therefore be correctly weighted to account for this, i.e. by introducing a covariance matrix: $\mathbf{\Sigma}^{-1}$, whose diagonal elements are the variances in the log-transformed data. To a first order approximation, this can be expressed as
$\sigma^2_{\ln I_i} = \sigma^2 / I_i^2$. The solution is therefore:

$$\mathbf{D} = \left(\mathbf{B}^T \Sigma^{-1} \mathbf{B}\right)^{-1} \mathbf{B} \Sigma^{-1} \mathbf{S} \qquad [5]$$

The diagonal elements of the first term in Eq. [5], $\left(\mathbf{B^T \Sigma^{-1} B}\right)^{-1}$, represent the predicted error variances of the estimated parameters. Under the assumption that $\mathbf{\Sigma}^{-1} = k\mathbf{I}$ (where k is a constant, and **I** is the identity matrix), this precision matrix is determined by the elements of **B** – which are under the experimenter's control. Further, Skare et al. (2000) showed, from Eq. [5], that the relative errors in the tensor matrix can be expressed in terms of the errors in the measurements, via the following expression

$$\frac{\|\Delta \mathbf{D}\|}{\|\mathbf{D}\|} \leq cond(\mathbf{B}) \frac{\|\Delta \mathbf{S}\|}{\mathbf{S}}, \qquad [6]$$

where *cond*(**B**) is the condition number of the matrix **B,** i.e.

$$cond(\mathbf{B}) = \sqrt{\frac{\lambda_{\max}\left(\mathbf{B}^{\mathrm{T}} \cdot \mathbf{B}\right)}{\lambda_{\min}\left(\mathbf{B}^{\mathrm{T}} \cdot \mathbf{B}\right)}}, \qquad [7]$$

where λ corresponds to the eigenvalues of the matrix in parentheses. It should be clear from Eq. [4] and Eq. [7] that for a fixed trace of the *b*-matrix, that the condition number will be solely dependent on the choice of gradient directions – which is something the operator has control over.

The conclusion is that the set of gradient directions chosen affects the errors in the estimated tensor elements. Skare et al. (2000) took the observation in Eq. [6] and used numerical optimization to find a set of gradient orientations that minimized the condition number. However, they found that results obtained with these schemes did not perform as well as schemes in which gradient encoding orientations were uniformly distributed in space (Jones et al. 1999). The reason for this was not immediately clear until Batchelor and colleagues (2003) showed that, for certain sampling schemes, the condition number depends on the relative orientation with respect to the laboratory frame of reference. In other words, the condition number is *rotationally invariant*. They did show, however, that schemes in which gradients are distributed as uniformly as possible in space (e.g., by pointing to the faces of an icosahedron), have rotationally *invariant* condition numbers. This introduces the concept of *statistical rotational invariance* – which means that the statistical properties of the tensor (e.g., variance / precision), will be independent of the orientation of the tensor.
A further observation made by both Skare et al. (2000) and Batchelor et al. (2003) is that, out of two schemes which have the same condition number but different numbers of *unique* gradient orientations, the scheme with more unique sampling orientations will outperform (in terms of statistical rotational invariance) the scheme with fewer directions. Jones (2004) systematically examined the effect of increasing the number of unique sampling orientations on statistical properties of the tensor and found that there is a clear advantage to using more than six directions – but that there

are diminishing returns after approximately 20 unique directions. There is ongoing debate, however, about optimal sampling orientations in the literature.

Parameters Derived From DT-MRI

There are three main parameters derived from DT-MRI: (a) Trace or Mean diffusivity; (b) Anisotropy; (c) Fibre Orientation. The first two can be computed from the eigenvalues and the latter from the eigenvectors.

Trace: Without doubt, the most clinically useful measure obtained from diffusion tensor imaging is the Trace. This is the sum of the three diagonal elements of the diffusion tensor (i.e. $D_{xx} + D_{yy} + D_{zz}$). The Trace/3 can be thought of as being equal to the *orientationally averaged* mean diffusivity. Note that, particularly in the earlier diffusion MRI literature, many alternative phrases have been used to describe this measure, including trace ADC and mean trace ADC. These terms are nonsensical since the trace is a property of tensors, while an ADC is a scalar quantity; the use of such terms should therefore be avoided. A remarkable property of the trace is that, in the range of diffusion weightings typically used in clinical studies (< 1000 s mm^{-2}), the mean diffusivity is fairly uniform throughout parenchyma (0.7×10^{-3} mm^2 s^{-1}). Although homogeneity makes it difficult to distinguish anatomical structures, it does offer the advantage that the effects of anisotropy do not confound detection of diffusion abnormalities, such as acute ischemic.

Anisotropy: Prior to the introduction of the tensor model into MRI by Basser *et al*., several indices for diffusivity were proposed, such as the ratio of ADCs obtained in two orthogonal directions. The limitation of such indices can be understood by imagining a set of fibers oriented at 45° to the *x*- and *y*-axes, the ratio ADC_y/ADC_x is equal to unity, for the fibers oriented along the *y*-axis, the ratio ADC_y/ADC_x takes its maximal value, and for the fibers oriented along the *x*-axis, the ratio takes its minimal value. As this measure depends on the orientation of the tissue with respect to the laboratory frame of reference, it is said to be *rotationally variant.*

Anisotropy indices formed from the eigenvalues of the tensor will, by definition, be rotationally invariant. The simplest anisotropy index, analogous to the ratio ADC_y/ADC_x would be the ratio of the largest to the smallest eigenvalue (i.e. λ_1/λ_3). However, as discussed later, it has been shown that sorting the eigenvalues according to their magnitude introduces a bias in the measurements at low signal-to-noise ratios (Pierpaoli *et al.* 1996). To circumvent this problem, indices that do not require sorting (Basser and Pierpaoli 1996; Pierpaoli and Basser 1996) have been proposed and have been shown to be less sensitive to the signal-to-noise ratio. The two most popular are the fractional anisotropy (*FA*) and relative anisotropy (*RA*), given by

$$FA = \sqrt{\frac{3}{2}} \frac{\sqrt{\left(\lambda_1 - \langle\lambda\rangle\right)^2 + \left(\lambda_2 - \langle\lambda\rangle\right)^2 + \left(\lambda_3 - \langle\lambda\rangle\right)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \qquad [8]$$

and

$$RA = \sqrt{\frac{1}{3}} \frac{\sqrt{\left(\lambda_1 - \langle\lambda\rangle\right)^2 + \left(\lambda_2 - \langle\lambda\rangle\right)^2 + \left(\lambda_3 - \langle\lambda\rangle\right)^2}}{\langle\lambda\rangle}, \qquad [9]$$

where

$$\langle \lambda \rangle = \frac{1}{3}(\lambda_1 + \lambda_2 + \lambda_3). \qquad [10]$$

The numerator for both terms is the same and is related to the variance of the three eigenvalues about their mean. The fractional anisotropy index normalizes this measure by the magnitude of the tensor as a whole. Just as the magnitude of a vector can be found from the sum of the squares of its individual components, the magnitude of the tensor is found from the sum of the squares of its eigenvalues. Thus, fractional anisotropy measures the fraction of the tensor that can be assigned to anisotropic diffusion. The fractional anisotropy index is appropriately normalized so that it takes values from zero (when diffusion is isotropic) to one (when diffusion is constrained along one axis only). The denominator of the relative anisotropy index is simply the mean diffusivity. This index is mathematically identical to a coefficient of variation, i.e. standard deviation divided by the mean. To ensure that this index scales from zero to one, Shimony *et al.* (1999) divided the relative anisotropy index by $\sqrt{2}$, and renamed the index A_σ, i.e.

$$A_\sigma = \frac{RA}{\sqrt{2}}, \qquad [11]$$

It should be noted that even though measures such as fractional anisotropy and relative anisotropy are less sensitive to noise than measures such as λ_1/λ_3, they are nevertheless sensitive to noise. As the signal-to-noise ratio is lowered, the anisotropy indices become increasingly overestimated (Pierpaoli *et al.* 1996). Thus comparisons of anisotropy indices obtained from different studies in which different imaging parameters have been used should be treated with caution.

Tensor Orientation: By finding the direction in which the motion of diffusing molecules is least hindered within each voxel, (given by the eigenvector associated with the largest eigenvalue), one can infer the dominant fiber orientation. By viewing fiber orientation in one voxel and following, by eye, a path of smooth transition in orientation from one voxel to the next, one can gain an impression of the trajectory of the major white matter pathways. This will be covered in detail in the lecture by Dr Andy Alexander.

Minimal Encoding Schemes

Given the three measures (trace, anisotropy and fiber orientation), it is interesting to pose the question whether one needs to collect all the data to estimate the tensor, if only one aspect is required. In terms of estimating the trace, $(D_{xx} + D_{yy} + D_{zz})$, it should be clear from Eq. [3] that if the pulse-sequence can be designed such that

$$b_{xy} = b_{yz} = b_{xz} = 0 \;\&\; b_{xx} = b_{yy} = b_{zz} = b, \qquad [12]$$

then the signal attenuation in Eq. [3] can be written as:

$$I = I_0 \exp\left(-b\left(D_{xx} + D_{yy} + D_{zz}\right)\right), \qquad [13]$$

Mori et al. (1995) and Wong et al. (1995) independently proposed schemes in which diffusion-encoding gradients were arranged on all three axes to satisfy the conditions given in Eq. [12], thereby producing a trace-weighted image, or (with just one additional image), a trace map very rapidly. Shrager and Basser (1998) addressed the question of how many diffusion-weighted images would be required to encode diffusion anisotropy. It can be shown that the minimum number if six – which, as

discussed above, is the minimum number required to estimate the tensor. Similar arguments can no doubt be applied to estimates of fiber orientation (i.e. one needs to estimate the diffusion tensor).

Further issues on optimisation (b-values, number of measurements made at each b-value, etc. etc.) will be discussed if time permits.

REFERENCES

Basser PJ, Matiello J, Le Bihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B* 1994; **103:** 247-254.

Basser PJ, Pierpaoli C. Microstructural and physiological features of tissue elucidated by quantitative-diffusion-tensor MRI. *J Magn Reson B* 1996; **111:** 209-219.

Batchelor PG, Atkinson D, Hill DL, Calamante F, Connelly A. Anisotropic noise propagation in diffusion tensor MRI sampling schemes. *Magn Reson Med*. 2003; **49:**1143-51.

Jones DK, Horsfield MA, Simmons A. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn. Reson. Med.* 1999; **42**: 515–525.

Jones DK. 2004.The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: A Monte Carlo study. *Magn Reson Med* **51:**807-815.

Mattiello J, Basser PJ, Le Bihan D. Analytical expression for the b matrix in NMR diffusion imaging and spectroscopy. *J. Magn.Reson. A* 1994; **108**: 131–141.

Mori, S., and van Zijl, P.C.M., Diffusion weighting by the trace of the diffusion tensor within a single scan, *Magn. Reson. Med*., 113, 265-270, 1995.

Pierpaoli C, Basser PJ. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med.* 1996; **36:** 893-906.

Pierpaoli C, Jezzard P, Basser PJ, Barnett AS. Diffusion tensor MR imaging of the human brain. *Radiology* 1996; **201:** 637-648.

Shrager RI, Basser PJ. Anisotropically weighted MRI. *Magn Reson Med* 1998; 40:160-165.

Skare S, Hedehus M, Moseley ME, Li TQ. Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. *J Magn Reson* 2000; **147:** 340-352.

Wong EC, Cox RW, Song AW. Optimized isotropic diffusion weighting. *Magn. Reson. Med.* 1995; **34**(2): 139–143.